Bayesian Inference and Decision Theory – A Coherent Framework for Decision Making in Natural Resource Management

Robert M. Dorazio

U. S. Geological Survey

Florida Caribbean Science Center

7920 NW 71 Street

Gainesville, Florida 32653

email: bdorazio@usgs.gov

Fred A. Johnson

U.S. Fish and Wildlife Service

Division of Migratory Bird Management

7920 NW 71 Street

Gainesville, Florida 32653

email: fred_a_johnson@fws.gov

Abstract

Bayesian inference and decision theory may be used in the solution of relatively complex problems of natural resource management, owing to recent advances in statistical theory and computing. In particular, Markov chain Monte Carlo algorithms provide a computational framework for fitting models of adequate complexity and for evaluating the expected consequences of alternative management actions. We illustrate these features using a hypothetical example based on management of waterfowl habitat.

Key words: Adaptive; Bayes theorem; Habitat; Optimal decision.

1 Introduction

Formal methods of decision making in natural resource management combine models of the dynamics of an ecological system with an objective function, which values the outcomes of alternative management actions. A common decision-making problem involves a temporal sequence of decisions, each alike in kind, but where the optimal action at each decision point may depend on time and/or system state (Possingham 1997). The goal of the manager, then, is to develop a decision rule (or management strategy) that prescribes management actions for each time or system state that are optimal with respect to the objective function. Examples of this kind of decision problem include direct manipulation of plant or animal populations through harvesting, stocking, or transplanting, as well as indirect population management through chemical or physical manipulation of relevant habitat attributes. Often, these problems also have a spatial aspect, wherein management decisions are required at different locations.

A rigorous analysis of such decision problems requires specification of (1) an objective function for evaluating alternative management strategies; (2) predictive models of system dynamics formulated in terms of quantities relevant to management objectives; (3) a finite set of alternative management actions, including any constraints on their use; and (4) a monitoring program to follow the system's evolution and responses to management. The objective function specifies the value of alternative management actions and usually accounts for both benefits and costs, as well as conditional constraints. The predictive models must be realistic enough to mimic the relevant behaviors of ecological systems, which often are complex (i.e., include many interacting components), are characterized by spatial, temporal, and organizational heterogeneity, and involve nonlinear dynamics. Thus, specification of an objective function and of useful system models can often be a demanding and difficult task in practical applications of decision theory to problems of natural resource management.

Perhaps even greater challenges are induced, however, by uncertainty in the predictions of management outcomes. This uncertainty may stem from incomplete control of management actions, errors in measurement and sampling of ecological systems, environmental variability, or incomplete knowledge of system behavior (Williams et al. 1996). A failure to recognize and account for these sources of uncertainty can severely depress management performance and, in some cases, has led to catastrophic environmental and economic losses (Ludwig et al. 1993). Accordingly, there has been a growing interest in the theory of stochastic decision processes, and in practical methods for deriving optimal (or at least robust) solutions (Walters and Hilborn 1978, Hilborn 1987, Williams 1989). Recently, there has been a particular emphasis on methods that can account for uncertainty in the dynamics of ecological systems and in their responses to both controlled and uncontrolled factors (Walters 1986). This uncertainty can be characterized by continuous or discrete distributions of model parameters (or by discrete distributions of alternative model forms), which are hypothesized or estimated from historical data (e.g., see Walters 1975, Johnson et al. 1997). In this manner, model uncertainty can be accommodated in solutions of decision problems in exactly the same manner as environmental variation and incomplete management control (Walters 1975). An important conceptual advance, however, has been the recognition that these probability distributions are not static, but rather evolve over time as new observations of system behaviors are accumulated during the management process (Walters 1986). The currently popular notion of adaptive resource management involves efforts that attempt to account for these dynamics of uncertainty in making management decisions (Walters 1986, Walters and Holling 1990, Williams 1996).

In this paper we argue that Bayesian inference and decision theory provide a coherent, theoretical framework for decision making in problems of natural resource management. Bayesian inference includes a probabilistic approach for sequentially updating beliefs (specified in terms of model parameters) as new information is acquired through monitoring and for predicting the consequences of future management actions, while properly accounting for uncertainty in the updated beliefs. In Bayesian decision theory, management objectives are specified as a function of model predictions (and/or parameters), and the expected consequences of any particular management action are calculated by integrating over the uncertainty in both model parameters and predictions.

The potential applicability of Bayesian methods in problems of natural resource management or conservation has been recognized previously (Ellison 1996, Bergerud and Reed 1998, Wade 2000); however, only recently have advances in statistical theory and computing allowed fairly complex, and hopefully more realistic, models to be fitted and used in decision making. Markov chain Monte Carlo algorithms (Gelfand and Smith 1990, Smith and Roberts 1993, Gilks et al. 1996), for example, are currently used in Bayesian analyses to fit complex models that were considered intractable only a decade ago.

In this paper we illustrate the Bayesian approach to inference and decision-making using a hypothetical example based on management of waterfowl habitat. This example is motivated by an actual problem and, although greatly simplified, includes several features that are common in problems of natural resource management. Our objective is to illustrate the general utility of the Bayesian approach in these problems, taking advantage of modern technological advances in Bayesian computation.

2 Inference and Decision-Making in a Problem of Habitat Management

2.1 Background and Setting

Suppose a moderately large property (say, on the order of a few thousand acres) is managed to provide habitat for waterfowl that may only be present during a brief overwintering period. (Migratory ducks that originate and live primarily in northeastern North America but migrate to Florida for the winter are good examples.) The wildlife managers responsible for this property believe that a combination of emergent vegetation interspersed with about 50% open water provides nearly ideal habitat for these waterfowl. Managers can regulate water levels on the property with reasonably good precision (owing to the presence of impoundments); however, there is considerable uncertainty about how to control growth of vegetation to provide suitable habitat. Various types of physical manipulation, such as burning, cutting, or grazing of vegetation, represent possible management actions for controlling growth; however, the effects of these manipulations are not well understood and are difficult to predict. Nonetheless, wildlife managers must develop a strategy that combines water-level regulation with one or more types of physical manipulation of the vegetation to achieve their objective of 50% open water and 50% vegetation cover.

Assume that the property to be managed is subdivided into n non-overlapping plots of approximately equal size and shape that can be manipulated in various ways to alter vegetation cover. Let $\boldsymbol{x}$ denote a $q \times 1$ design vector that specifies which of the q possible management actions (i.e., manipulations) is applied to an individual plot. Without loss of generality, we define $\boldsymbol{x}$ using a "centered" parameterization wherein $\boldsymbol{x} = (1, 0, \ldots, 0)^T$ specifies the first management action, $\boldsymbol{x} = (0, 1, \ldots, 0)^T$ specifies the second management action, and so on. The exponent, T, indicates the transpose of a matrix or vector.

In the first year of management suppose we have a procedure (e.g., randomization) for deciding which of the q possible management actions is applied to each of the n plots. In other words, we have a way of assigning a value to $\boldsymbol{x}_{i1}$, the design vector for the ith plot ($i = 1, \ldots, n$) at time $t = 1$. For the moment, we assume that each management action can be applied without error (i.e., uncertainty due to partial controllability of management actions is negligible). Our initial management actions may be summarized in a $n \times q$ matrix $\boldsymbol{X}_1 = (\boldsymbol{x}_{11}, \boldsymbol{x}_{21}, \ldots, \boldsymbol{x}_{n1})^T$. Suppose we have implemented these actions and observed the vegetation cover in each plot. Denoting these n responses in vegetation cover with an $n \times 1$ vector $\boldsymbol{y}_1$ (subscript indicates $t = 1$), we summarize the results of the first year of management actions as follows:

Plot	ManagementAction	VegetationCover
1	$\boldsymbol{x}_{11} = (1,0)^T \Rightarrow$ burning	y_{11}
2	$\boldsymbol{x}_{21} = (1,0)^T \Rightarrow$ burning	y_{21}
3	$\boldsymbol{x}_{31} = (0,1)^T \Rightarrow$ cutting	y_{31}
$\vdots$	$\vdots$	$\vdots$
n	$\boldsymbol{x}_{n1} = (0,1)^T \Rightarrow$ cutting	y_{n1}

(Only $q = 2$ management actions are illustrated for ease of presentation.)

Given these results, we now require a procedure for selecting a new set of management actions to be implemented in the second year. Our selection should depend on the plot-specific responses of vegetation to management actions applied in the previous year and on the need to satisfy the overall management objective of 50% vegetation cover. In other words, we need a procedure that specifies $\boldsymbol{X}_2$, the design at $t = 2$, given our management objective and our current beliefs.

2.2 Modeling Consequences of Management Actions

Statistical models provide an essential framework for specifying our beliefs and for making evidentiary conclusions or predictions based on those beliefs and on the available data. In our habitat-management problem, a statistical model is needed to provide a quantitative, unambiguous description of the processes thought to be responsible for producing plot-specific differences in vegetation cover. The model allows us to infer which processes are most important in terms of well-defined criteria (i.e., model parameters) and to predict the consequences of future management actions given our current level of understanding and current estimates of uncertainty.

We consider the following, relatively simple model of plot-specific vegetation responses over a period of τ years. Assume the vegetation cover in an individual plot depends on both the current type of management and on past levels of vegetation observed in that plot. We can specify these dependencies using a first-order, autoregressive model:

$$\begin{pmatrix} Y_{i1} \\ Y_{i2} \\ Y_{i3} \\ \vdots \\ Y_{i\tau} \end{pmatrix} \sim \mathrm{N}\left(\begin{pmatrix} \boldsymbol{x}_{i1}^T\boldsymbol{\beta} \\ \boldsymbol{x}_{i2}^T\boldsymbol{\beta} \\ \boldsymbol{x}_{i3}^T\boldsymbol{\beta} \\ \vdots \\ \boldsymbol{x}_{i\tau}^T\boldsymbol{\beta} \end{pmatrix}, \ \frac{\sigma^2}{(1-\rho^2)} \begin{pmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{\tau-1} \\ \rho & 1 & \rho & & \\ \rho^2 & \rho & 1 & & \\ \vdots & & & \ddots & \\ \rho^{\tau-1} & & & & 1 \end{pmatrix} \right) \tag{1}$$

where Y_{it} is a random variable for vegetation cover in plot i in year t, $\boldsymbol{x}_{it}$ specifies the management action applied to plot i in year t, and $\boldsymbol{\beta}$, σ^2, and ρ are model parameters. Given the "centered" parameterization implied in our definition of $\boldsymbol{x}$, each element of $\boldsymbol{\beta}$ corresponds to the mean vegetation cover associated with a distinct management action. The parameter ρ denotes the correlation between vegetation responses observed in consecutive years.

For our purposes, it is useful to express the plot-specific temporal dependence in vegetation cover in the following form, which is equivalent to (1):

$$(Y_{it} \mid \boldsymbol{x}_{it}, \boldsymbol{\beta}, \sigma^2, \rho, \, y_{i,t-1}, \, \boldsymbol{x}_{i,t-1}) \sim \begin{cases} \mathrm{N}(\boldsymbol{x}_{it}^T\boldsymbol{\beta}, \ \sigma^2/(1-\rho^2)) & \text{if } t = 1, \\ \mathrm{N}\left(\boldsymbol{x}_{it}^T\boldsymbol{\beta} + \rho\,(y_{i,t-1} - \boldsymbol{x}_{i,t-1}^T\boldsymbol{\beta}), \ \sigma^2\right) & \text{if } t > 1. \end{cases} \tag{2}$$

Thus, by conditioning on the sequence of past observations $(y_{i1}, \ldots, y_{i,t-1})$, we express the conditional mean of the ith plot's vegetation cover in year $t\,(> 1)$ in terms of present and past management actions ($\boldsymbol{x}_{it}$ and $\boldsymbol{x}_{i,t-1}$, respectively). This form of conditioning induces a temporal dynamic that has important implications for the adaptive selection of plot-specific management actions and will be discussed more fully in Section 2.4.

So far we have considered only how vegetation cover might respond to changes in management actions within a single plot. All plots on the property are monitored and manipulated in an adaptive approach to management; therefore, we require a model of the vegetation responses in all plots. The simplest assumption to consider is that plot-specific responses are conditionally independent; thus, their joint density is

$$f(\boldsymbol{y}_t \mid \boldsymbol{X}_t, \boldsymbol{\theta},\, \boldsymbol{y}_{t-1}, \boldsymbol{X}_{t-1}) = \prod_{i=1}^{n} f(y_{it} \mid \boldsymbol{x}_{it}, \boldsymbol{\theta},\, y_{i,t-1}, \boldsymbol{x}_{i,t-1}), \tag{3}$$

where $f(y_{it} \mid \boldsymbol{x}_{it}, \boldsymbol{\theta},\, y_{i,t-1}, \boldsymbol{x}_{i,t-1})$ specifies the conditional distribution (in (2)) of vegetation cover in the ith plot and $\boldsymbol{\theta} = (\boldsymbol{\beta}, \rho, \sigma^2)^T$ is a vector of model parameters. The right-hand-side of (3) would be more complicated if we had assumed that vegetation cover depended, in part, on the proximity of one plot to another. For now, however, we ignore spatial dependence in vegetation cover (but see Section 3).

2.3 Bayesian Updating of Model Parameters

Armed with a model of the responses in vegetation to different types of physical manipulation, we now describe how Bayesian inference may be used to update our beliefs about model parameters as new management actions are implemented over time. At the end of the first year, we have implemented a set of management actions ($\boldsymbol{X}_1$) and observed the responses of vegetation cover to those management actions ($\boldsymbol{y}_1$). Applying Bayes Theorem yields the posterior distribution of model parameters,

$$p(\boldsymbol{\beta}, \sigma^2, \rho \mid \boldsymbol{y}_1, \boldsymbol{X}_1) = \frac{f(\boldsymbol{y}_1 \mid \boldsymbol{X}_1, \boldsymbol{\beta}, \sigma^2, \rho)\ \pi(\boldsymbol{\beta}, \sigma^2, \rho)}{\int f(\boldsymbol{y}_1 \mid \boldsymbol{X}_1, \boldsymbol{\theta})\ \pi(\boldsymbol{\theta})\, d\boldsymbol{\theta}}, \tag{4}$$

which indicates how our initial opinion of the model parameters (specified in the prior distribution $\pi(\boldsymbol{\beta}, \sigma^2, \rho)$) is modified in light of the observed responses of vegetation to management. The contribution of these data to the posterior is called the likelihood function, which we denote by f. Since one year of data provides no information about the temporal dependence of vegetation cover within each plot, information about ρ in the posterior for $t = 1$ will be identical to that specified in the prior $\pi(\boldsymbol{\beta}, \sigma^2, \rho)$. Our opinions about $\boldsymbol{\beta}$ and σ^2, on the other hand, are likely to be influenced by the first year's results.

Now imagine that we have selected and implemented a set of management actions in the second year ($\boldsymbol{X}_2$) and observed the responses in vegetation cover ($\boldsymbol{y}_2$). Again, applying

Bayes Theorem yields the posterior distribution of model parameters

$$p(\boldsymbol{\beta}, \sigma^2, \rho \mid \boldsymbol{y}_1, \boldsymbol{y}_2, \boldsymbol{X}_1, \boldsymbol{X}_2) = \frac{f(\boldsymbol{y}_2 \mid \boldsymbol{X}_2, \boldsymbol{\beta}, \sigma^2, \rho,\ \boldsymbol{y}_1, \boldsymbol{X}_1)\ p(\boldsymbol{\beta}, \sigma^2, \rho \mid \boldsymbol{y}_1, \boldsymbol{X}_1)}{\int f(\boldsymbol{y}_2 \mid \boldsymbol{X}_2, \boldsymbol{\theta},\ \boldsymbol{y}_1, \boldsymbol{X}_1)\ p(\boldsymbol{\theta} \mid \boldsymbol{y}_1, \boldsymbol{X}_1)\, d\boldsymbol{\theta}}, \tag{5}$$

which reveals how our opinion of the model parameters at the end of the first year is modified by the results observed in the second year. In particular, ρ may now be updated based on the second year of responses in vegetation cover.

Using successive applications of Bayes Theorem, it is easy to show that the posterior distribution of model parameters at the end of the tth year is

$$p(\boldsymbol{\theta} \mid \boldsymbol{y}_1, \ldots, \boldsymbol{y}_t, \boldsymbol{X}_1, \ldots, \boldsymbol{X}_t) = \frac{f(\boldsymbol{y}_t \mid \boldsymbol{X}_t, \boldsymbol{\theta},\ \boldsymbol{y}_{t-1}, \boldsymbol{X}_{t-1})\ p(\boldsymbol{\theta} \mid \boldsymbol{y}_1, \ldots, \boldsymbol{y}_{t-1}, \boldsymbol{X}_1, \ldots, \boldsymbol{X}_{t-1})}{\int f(\boldsymbol{y}_t \mid \boldsymbol{X}_t, \boldsymbol{\psi},\ \boldsymbol{y}_{t-1}, \boldsymbol{X}_{t-1})\ p(\boldsymbol{\psi} \mid \boldsymbol{y}_1, \ldots, \boldsymbol{y}_{t-1}, \boldsymbol{X}_1, \ldots, \boldsymbol{X}_{t-1})\, d\boldsymbol{\psi}}, \tag{6}$$

where $\boldsymbol{\psi}$ represents all possible values of the model parameters. Thus, Bayes Theorem provides a general method for sequentially updating our beliefs and quantifying our uncertainty about model parameters as new results are acquired. In Section 2.4 we describe how Bayesian updating is used to evaluate the consequences of future management actions and thereby help to achieve the overall management objective of 50% vegetation cover.

2.4 Computing an Optimal Set of Management Actions

Our overall management objective (50% vegetation cover) is defined in terms of quantities that are directly observable, unlike the unobservable model parameters. To evaluate the consequences of future management actions, we therefore require predictions of the (observable) vegetation cover in each plot, given what we have learned from past observations.

Let $\tilde{\boldsymbol{y}}_t$ denote an $n \times 1$ vector of plot-specific predictions of vegetation cover in year t. The posterior predictive distribution of $\tilde{\boldsymbol{y}}_t$

$$p(\tilde{\boldsymbol{y}}_t \mid \tilde{\boldsymbol{X}}_t,\ \boldsymbol{y}_1, \ldots, \boldsymbol{y}_{t-1}, \boldsymbol{X}_1, \ldots, \boldsymbol{X}_{t-1}) = \int f(\tilde{\boldsymbol{y}}_t \mid \tilde{\boldsymbol{X}}_t, \boldsymbol{\theta},\ \boldsymbol{y}_{t-1}, \boldsymbol{X}_{t-1})\ p(\boldsymbol{\theta} \mid \boldsymbol{y}_1, \ldots, \boldsymbol{y}_{t-1}, \boldsymbol{X}_1, \ldots, \boldsymbol{X}_{t-1})\, d\boldsymbol{\theta} \tag{7}$$

specifies our uncertainty in predictions of vegetation cover in year t, given a proposed set of management actions ($\tilde{\boldsymbol{X}}_t$) and the sequence of vegetation covers ($\boldsymbol{y}_1, \ldots, \boldsymbol{y}_{t-1}$) observed after implementation of management actions ($\boldsymbol{X}_1, \ldots, \boldsymbol{X}_{t-1}$) in years 1 through $t-1$. The posterior predictive distribution properly accounts for all sources of uncertainty because it integrates the conditional likelihood of plot-specific predictions of vegetation cover over the posterior uncertainty of all model parameters.

We now describe how (7) is used to select future management actions that maximize our opportunity to achieve the overall management objective of 50% vegetation cover. We denote these management actions as "optimal." Assume that a set of plot-specific management actions $\boldsymbol{X}_1$ has been implemented and that the vegetation responses to those actions $\boldsymbol{y}_1$ have been observed. We require a procedure for selecting an optimal set of management actions to be implemented in year 2. Let $l(\tilde{\boldsymbol{y}}_2, c)$ denote a function that specifies the scalar-valued loss incurred when our predictions of vegetation cover differ from the target value ($c = 50\%$). For example, $l(\tilde{\boldsymbol{y}}_2, c) = \sum_{i=1}^{n} |\tilde{y}_{i2} - c|$ is an absolute-error loss function, which equals the sum of the absolute discrepancies between plot-specific predictions of vegetation cover and the target value.

The loss function $l(\tilde{\boldsymbol{y}}_2, c)$ allows us to develop an unambiguous, mathematical description of our overall management objective. Specifically, we seek a (future) management action $\tilde{\boldsymbol{X}}_2$ that minimizes the loss that can be expected given the posterior uncertainty in plot-specific predictions of vegetation cover. We denote this expected loss by

$$\begin{aligned} \bar{l}(\tilde{\boldsymbol{X}}_2 \mid \boldsymbol{y}_1, \boldsymbol{X}_1) &= \mathrm{E}_{(\tilde{\boldsymbol{y}}_2 | \tilde{\boldsymbol{X}}_2, \boldsymbol{y}_1, \boldsymbol{X}_1)} \left[l(\tilde{\boldsymbol{y}}_2, c) \right] \\ &= \int l(\tilde{\boldsymbol{y}}_2, c)\; p(\tilde{\boldsymbol{y}}_2 \mid \tilde{\boldsymbol{X}}_2, \boldsymbol{y}_1, \boldsymbol{X}_1)\, d\tilde{\boldsymbol{y}}_2, \end{aligned} \tag{8}$$

which reveals the crucial role of the posterior predictive distribution $p(\tilde{\boldsymbol{y}}_2 \mid \tilde{\boldsymbol{X}}_2, \boldsymbol{y}_1, \boldsymbol{X}_1)$ in this problem. Our overall management objective may now be stated succinctly: Find an optimal set of future management actions $\tilde{\boldsymbol{X}}_2^*$ such that

$$\tilde{\boldsymbol{X}}_2^* = \arg \min_{\tilde{\boldsymbol{X}}_2} \left[\bar{l}(\tilde{\boldsymbol{X}}_2 \mid \boldsymbol{y}_1, \boldsymbol{X}_1) \right]. \tag{9}$$

The computations involved in solving (9) may be formidable; however, in principle a solution can always be found, assuming that one exists (see Section 2.5.1 for an example where no optimum exists). We have assumed that one of q management actions can be implemented in each of the n plots; therefore, there are q^n possible values of $\tilde{\boldsymbol{X}}_2$ to compare in the search for an optimal set of management actions $\tilde{\boldsymbol{X}}_2^*$.

Although we cannot seriously expect our model assumptions to remain valid indefinitely long, we can also compute an optimal *sequence* of future management actions. Suppose we have observed $\boldsymbol{X}_1$ and $\boldsymbol{y}_1$ and want to predict an optimal sequence of future management actions $(\tilde{\boldsymbol{X}}_2^*, \tilde{\boldsymbol{X}}_3^*, \ldots, \tilde{\boldsymbol{X}}_\tau^*)$ to be implemented in the next $\tau - 1$ years. Let $l(\tilde{\boldsymbol{y}}_2, \ldots, \tilde{\boldsymbol{y}}_\tau, c)$ denote a scalar-valued loss function that specifies the loss incurred when future predictions of vegetation cover fail to meet the objective of $c = 50\%$. As in (8), we define the expected loss through year τ as follows:

$$\begin{aligned}\bar{l}(\tilde{\boldsymbol{X}}_2, \ldots, \tilde{\boldsymbol{X}}_\tau \mid \boldsymbol{y}_1, \boldsymbol{X}_1) &= \mathrm{E}_{(\tilde{\boldsymbol{y}}_2,\ldots,\tilde{\boldsymbol{y}}_\tau \mid \tilde{\boldsymbol{X}}_2,\ldots,\tilde{\boldsymbol{X}}_\tau, \boldsymbol{y}_1, \boldsymbol{X}_1)} [l(\tilde{\boldsymbol{y}}_2, \ldots, \tilde{\boldsymbol{y}}_\tau, c)] \qquad (10)\\ &= \int l(\tilde{\boldsymbol{y}}_2, \ldots, \tilde{\boldsymbol{y}}_\tau, c)\, p(\tilde{\boldsymbol{y}}_2, \ldots, \tilde{\boldsymbol{y}}_\tau \mid \tilde{\boldsymbol{X}}_2, \ldots, \tilde{\boldsymbol{X}}_\tau, \boldsymbol{y}_1, \boldsymbol{X}_1)\, d\tilde{\boldsymbol{y}}\end{aligned}$$

where $\tilde{\boldsymbol{y}} = (\tilde{\boldsymbol{y}}_2, \ldots, \tilde{\boldsymbol{y}}_\tau)^T$. The solution to our problem is the sequence of future management actions $(\tilde{\boldsymbol{X}}_2, \ldots, \tilde{\boldsymbol{X}}_\tau)$ that minimizes (10). Although numerical evaluations of (10) will be computationally expensive, they are feasible. For example, our model implies that a random draw from the posterior predictive distribution $(\tilde{\boldsymbol{y}}_2, \ldots, \tilde{\boldsymbol{y}}_\tau \mid \tilde{\boldsymbol{X}}_2, \ldots, \tilde{\boldsymbol{X}}_\tau, \boldsymbol{y}_1, \boldsymbol{X}_1)$ can be obtained by computing random draws from an appropriately ordered sequence of conditional posterior predictive distributions since

$$\begin{aligned}p(\tilde{\boldsymbol{y}}_2, \ldots, \tilde{\boldsymbol{y}}_\tau \mid \tilde{\boldsymbol{X}}_2, \ldots, \tilde{\boldsymbol{X}}_\tau, \boldsymbol{y}_1, \boldsymbol{X}_1) = p(\tilde{\boldsymbol{y}}_2 \mid \tilde{\boldsymbol{X}}_2, \boldsymbol{y}_1, \boldsymbol{X}_1)\, p(\tilde{\boldsymbol{y}}_3 \mid \tilde{\boldsymbol{y}}_2, \tilde{\boldsymbol{X}}_2, \tilde{\boldsymbol{X}}_3, \boldsymbol{y}_1, \boldsymbol{X}_1)\\ \cdots\; p(\tilde{\boldsymbol{y}}_\tau \mid \tilde{\boldsymbol{y}}_2, \ldots, \tilde{\boldsymbol{y}}_{\tau-1}, \tilde{\boldsymbol{X}}_2, \ldots, \tilde{\boldsymbol{X}}_\tau, \boldsymbol{y}_1, \boldsymbol{X}_1). \qquad (11)\end{aligned}$$

2.5 Numerical Examples

In this section 3 hypothetical data sets are used to clarify by example how data may be used to inform management decisions. The data are analyzed using the autoregressive model developed in Section 2.2, and decisions are made using the framework described in Sections 2.3 and 2.4. Computational details about sampling from posterior and posterior-predictive distributions are described in the appendix.

2.5.1 Equivocal responses in vegetation cover

In the first year of management suppose one of two types of management actions (denoted by $X_1 = 1$ and $X_1 = 2$) are randomly assigned to each of 4 plots. After doing so, we observe the vegetation cover (y_1, as a proportion) in each plot as follows:

Plot	X_1	y_1
1	1	0.15
2	2	0.55
3	2	0.85
4	1	0.45

The sample mean vegetation covers associated with management actions 1 and 2 (0.30 and 0.70, respectively) are equidistant from $c = 0.50$, the level of vegetation cover specified as our management objective. What plot-specific management actions $\tilde{\boldsymbol{X}}_2$ should be taken in year 2, given the vegetation responses observed year 1?

First we specify the management objective by assuming an absolute-error loss function, $l(\tilde{\boldsymbol{y}}_2, 0.50) = \sum_i |\tilde{y}_{i2} - 0.50|$, which quantifies the total discrepancy between predicted plot-specific vegetation cover and $c = 0.50$. The optimal set of management actions includes those which minimize the expected loss, averaging over the posterior uncertainty in plot-specific predictions of vegetation cover (as in (8)). In this case there are 16 ($= 2^4$) possible combinations of management actions to be compared (indicated in the columns below):

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	1	2	2	1	1	1	1	1	1	2	2	2	2	2	2
1	2	1	2	1	1	1	2	2	2	1	1	1	2	2	2
1	1	1	1	1	2	2	1	2	2	1	2	2	1	2	2
1	1	1	1	2	1	2	2	1	2	2	1	2	2	1	2

To complete a Bayesian analysis of the data from year 1, we assume mutually independent, Uniform(0,1) prior distributions for each component of $\boldsymbol{\beta}$ (the treatment-effect parameters), a conjugate Inverse-Gamma(0.1,0.1) prior for σ^2, and a fixed value for ρ (the parameter for temporal dependence within plots). Assuming ρ to be fixed is equivalent to specifying a prior for ρ with a point mass of 1.0 at the fixed value. It is necessary to specify a strongly informative prior for ρ because data from only 1 year provide no information about temporal dependence within plots. However, regardless of the value selected for ρ, the posterior means for β_1 and β_2 are 0.36 and 0.64, respectively, which reflects a Bayesian compromise (sometimes called "shrinkage") between the prior means (0.50 and 0.50) and the sample means (0.30 and 0.70).

Generally speaking, the optimal set of management actions proposed for year 2 depends on the value assumed for ρ. If we assume $\rho = 0$ (i.e., plot-specific responses in years 1 and 2 are uncorrelated), the expected losses of all 16 experimental designs are approximately equal (expected loss = 1.59, Monte Carlo standard error = 0.01). In other words, if no time dependence in vegetation cover is assumed, then any set of management actions is as good as any other, and the new set of management actions can be selected randomly and still be optimal (i.e., still be consistent with the management objective). This decision seems quite sensible because the posterior mean vegetation responses, being equidistant from $c = 0.50$, do not favor any of the 16 possible sets of management actions.

2.5.2 Favored responses in vegetation cover

This example is identical to the previous one except that we observe a different set of vegetation responses at the end of the first year:

Plot	X_1	y_1
1	1	0.15
2	2	0.55
3	2	0.65
4	1	0.25

Intuitively, one might guess that management action 2 is favored for selection in year 2 because the sample mean vegetation response to action 2 (0.60) is closer to the management objective of $c = 0.50$ than the mean vegetation response to management action 1 (0.20). This, in fact, turns out to be correct. If we assume that $\rho = 0$ (as in the previous example), the expected losses of the 16 experimental designs are

1	2	3	4	5	6	7	8
1.55	1.49	1.49	1.46	1.50	1.51	1.47	1.45
9	10	11	12	13	14	15	16
1.46	1.41	1.46	1.46	1.41	1.41	1.41	1.36

All 4 plots receive management action 2 under the optimal design (#16) with the lowest expected loss (1.36).

2.5.3 Equivocal and correlated responses in vegetation cover

The equivocal vegetation responses assumed in Section 2.5.1 failed to favor selection of any of the 16 possible management actions when plot-specific responses in years 1 and 2 were assumed to be uncorrelated. Based on this analysis, suppose we decide to leave the design unchanged in year 2 and observe an additional year of vegetation responses:

Plot	X_1	y_1	X_2	y_2
1	1	0.15	1	0.25
2	2	0.55	2	0.50
3	2	0.85	2	0.75
4	1	0.45	1	0.50

Notice that the sample mean vegetation covers associated with management actions 1 and 2 (0.3375 and 0.6625, respectively) are still equidistant from our management objective of $c = 0.50$. However, now there is enough information in the data to estimate the interannual dependence (or correlation) in plot-specific vegetation responses; therefore, we can examine the influence of ρ in the selection of management actions proposed for year 3.

Suppose we assume a mutually independent, Uniform(0,1) prior for each component of $\boldsymbol{\beta}$, a conjugate Inverse-Gamma(0.1,0.1) prior for σ^2, and a Uniform(-1,1) prior for ρ. A Bayesian analysis of the observed data yields a posterior mean for $\boldsymbol{\beta}$ that approximately equals the sample mean (0.35 and 0.65, respectively (Figure 1)). The posterior distribution of ρ (Figure 1) is highly skewed (mean = 0.37, median = 0.43) and indicates that the vegetation responses within each plot are positively correlated.

The expected losses used to compare different sets of management actions proposed for year 3 depend, as usual, on the posterior predictions of vegetation responses to those actions. However, given the positive interannual correlation in plot-specific responses that we have estimated from years 1 and 2, we do not expect all designs to have the same expected loss, even though the posterior mean vegetation responses to the 2 management actions are equidistant from $c = 0.50$. In fact, the expected losses are quite different, as indicated in the following table:

1	2	3	4	5	6	7	8
1.077	1.008	1.069	1.005	1.143	1.077	1.151	1.077
9	10	11	12	13	14	15	16
1.007	1.077	1.148	1.072	1.147	1.071	1.000	1.074

Design #15 provides the optimal set of management actions because it has the lowest expected loss (1.000). However, designs #2, #4, and #9 provide almost the same expected loss given the Monte Carlo standard error (0.005) of the estimates. In design #15 only the first plot receives a change in management compared to the previous 2 years. The predicted mean vegetation cover of plot 1 under design #15 is 0.60, which is closer to the management objective of $c = 0.5$ than the mean vegetation cover of 0.31, which is predicted if the management actions in plot 1 are left unchanged (design #9). This example illustrates that the temporal dependence in vegetation responses can exert considerable influence in the selection of alternative management actions.

3 Discussion

Management of natural resources generally involves a repeating sequence of data collection (monitoring), assessment (analysis of data and prediction of consequences of proposed management actions), and implementation (actions or manipulations intended to achieve management objectives). This sequence essentially represents an iterative updating of beliefs that includes learning from data and making decisions in the presence of uncertainty, activities which are inherent features of the Bayesian paradigm.

We have demonstrated that Bayesian inference and decision theory may be used in the solution of relatively complex problems of natural resource management, owing to recent advances in statistical computing. Our hypothetical problem of habitat management (Section 2), though greatly simplified, includes several common features of actual problems of natural resource management. For example, we assumed that changes in system state (plot-specific vegetation) depend on proposed and past management actions and on the past state of the system. State-dependent dynamics are often justified on scientific (problem-specific) grounds; however, they are sensible also in cases where the proximate causes of state dependence are poorly understood (and unobserved) but necessary for accurate predictions of future system state.

Actual problems of natural resource management often contain additional features that add complexity to models of system dynamics or to the loss functions used in specifying management objectives. Modern Bayesian methods of inference and decision-making are capable of accommodating many, if not all, of these additional complexities. For example, system dynamics frequently are influenced by factors that cannot be controlled by managers. Uncertainties in system responses to management actions may be induced by environmental variability or by errors in sampling, measurement, or application of management actions. Alternatively, the sources of uncertainty may be difficult to identify and yet produce conspicuous patterns of variation in system responses (e.g., spatial correlations). A proper accounting of these additional sources of uncertainty requires modeling; however, if models are to be useful and relevant in decision-making, the models must include parameters that can be updated as new information is acquired through monitoring. The Bayesian paradigm provides a coherent framework for updating *any* of the parameters in a model of system dynamics, including ancillary parameters that do not represent the direct effects of management actions on system responses. In addition, there are virtually no limits to the complexity of models that can be entertained. Technological advancements in Bayesian computation currently permit sophisticated, hierarchical models of spatial and temporal dependence to be fitted with relative ease (Wikle et al. 1998, Datta et al. 2000).

In some problems of natural resource management, scientific reasoning may indicate that 2 or more structurally distinct models of system dynamics could be fitted to the data and used in decision making. In other problems the process of model selection may be somewhat arbitrary, and several models may fit the data equally well and provide plausible descriptions of the observations. In either case, it would seem more appropriate to predict the consequences of management actions by integrating over the posterior uncertainty of all models under consideration rather than by conditioning on the predictions of a single model. The Bayesian paradigm provides a straightforward method for averaging over model uncertainty (Draper 1995, Hoeting et al. 1999) that follows naturally from the calculus of probabilities and requires no additional theory or principles. Thus, it is entirely feasible to incorporate model uncertainty into the selection of alternative management actions.

In many problems of natural resource management, objectives are specified in terms of the cumulative losses and benefits obtained from a future sequence of management actions. The accumulated harvest of exploited fish or wildlife populations over some time frame is a good example. In such problems the expected loss used to evaluate alternative sequences of management actions generally depends on the joint distribution of predicted system responses (as in (10) for example); however, such loss functions pose no real difficulty for the Bayesian decision-making framework. Complicated loss functions also may occur in problems where a sequence of decisions is required but the relative effects of different management actions are poorly understood. In these problems managers initially may place greater value on learning about the magnitude of these effects than on achieving a particular management objective (e.g., a target level of vegetation cover). The rationale here is that learning may yield long-term benefits which exceed the short-term rewards that may be attained without an improved understanding of the effects of alternative management actions. Walters and Hilborn (1978) refer to these as dual-control problems that require "active adaptive" management. The competing objectives of dual-control problems must be specified in the loss function, which quantifies the benefits of learning from a proposed set of management actions. In a Bayesian treatment of the problem these benefits may be formulated in terms of the average discrepancy between the posterior distribution of model parameters and updates of the posterior that are predicted from the distribution of outcomes associated with a proposed set of management actions. Therefore, in dual-control problems loss functions will generally include model parameters (to quantify learning) *and* model predictions of observable system features (to quantify specific management objectives).

In this paper we have argued that modern methods of Bayesian inference and decision making are capable of solving complex problems of natural resource management. We anticipate widespread use of these methods in the near future, particularly as computing software is developed for estimating posterior distributions of model parameters and predictions (e.g., see the software guide in Appendix C of Carlin and Louis 2000).

Literature Cited

Bergerud, W. A., and Reed, W. J. 1998. Bayesian statistical methods. In Statistical methods for adaptive management studies, edited by V. Sit and B. Taylor, Land Management Handbook No. 42. Crown Publications, Victoria, British Columbia, pp. 89–104.

Carlin, B. P., and Louis, T. A. 2000. Bayes and empirical Bayes methods for data analysis, second edition. Chapman and Hall, Boca Raton, Florida.

Datta, G., Ghosh, M., and Waller, L. A. 2000. Hierarchical and empirical Bayes methods for environmental risk assessment. In Handbook of Environmental Statistics, edited by C. R. Rao and P. K. Sen. North Holland, pp. 223–245.

Draper, D. 1995. Assessment and propagation of model uncertainty (with discussion). Journal of the Royal Statistical Society, Series B **57**: 45–97.

Ellison, A. M. 1996. An introduction to bayesian inference for ecological research and environmental decision-making. Ecological Applications **6**: 1036–1046.

Gelfand, A. E., and Smith, A. F. M. 1990. Sampling-based approaches to calculating marginal densities. Journal of the American Statistical Association **85**: 398–409.

Gilks, W. R., Richardson, S., and Spiegelhalter, D. J. 1996. Markov Chain Monte Carlo in Practice. Chapman and Hall, London.

Gilks, W. R., and Wild, P. 1992. Adaptive rejection sampling for Gibbs sampling. Applied Statistics **41**: 337–348.

Gilks, W. R. 1992. Derivative-free adaptive rejection sampling for Gibbs sampling. In Bayesian Statistics 4, edited by J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith. Oxford University Press, Oxford, pp. 641–649.

Hilborn, R. 1987. Living with uncertainty in resource management. North American Journal of Fisheries Management **7**: 1–5.

Hoeting, J. A., Madigan, D., Raftery, A. E., and Volinsky, C. T. 1999. Bayesian model averaging: a tutorial (with discussion). Statistical Science **14**: 382–417.

Johnson, F. A., Moore, C. T., Kendall, W. L., Dubovsky, J. A., Caithamer, D. F., Kelley, J. R., Jr., and Williams, B. K. 1997. Uncertainty and the management of mallard harvests. Journal of Wildlife Management **61**: 202–216.

Ludwig, D., Hilborn, R., and Walters, C. 1993. Uncertainty, resource exploitation, and conservation: lessons from history. Science **260**: 17,36.

Possingham, H. P. 1997. State-dependent decision analysis for conservation biology. In The Ecological Basis of Conservation: Heterogeneity, Ecosystems, and Biodiversity, edited by S. T. A. Pickett, R. S. Ostfeld, M. Shachak, and G. E. Likens. pp. 298–304.

Smith, A. F. M., and Roberts, G. O. 1993. Bayesian computation via the Gibbs sampler and related Markov chain Monte Carlo methods (with discussion). Journal of the Royal Statistical Society, Series B **55**: 3–23.

Tanner, M. A. 1996. Tools for statistical inference: methods for the exploration of posterior distributions and likelihood functions, third edition. Springer-Verlag, New York.

Wade, P. R. 2000. Bayesian methods in conservation biology. Conservation Biology **14**: 1308–1316.

Walters, C. J., and Hilborn, R. 1978. Ecological optimization and adaptive management. Annual Review of Ecology and Systematics **9**: 157–188.

Walters, C. J., and Holling, C. S. 1990. Large-scale management experiments and learning by doing. Ecology **71**: 2060–2068.

Walters, C. J. 1975. Optimal harvest strategies for salmon in relation to environmental variability and uncertain production parameters. Journal of the Fisheries Research Board of Canada **32**: 1777–1784.

Walters, C. J. 1986. Adaptive Management of Renewable Resources. Macmillian, New York.

Wikle, C. K., Berliner, L. M., and Cressie, N. 1998. Hierarchical Bayesian space-time models. Environmental and Ecological Statistics **5**: 117–154.

Williams, B. K., Johnson, F. A., and Wilkins, K. 1996. Uncertainty and the adaptive management of waterfowl harvests. Journal of Wildlife Management **60**: 223–232.

Williams, B. K. 1989. Review of dynamic optimization methods in renewable natural resource management. Natural Resource Modeling **3**: 137–216.

Williams, B. K. 1996. Adaptive optimization and the harvest of biological populations. Mathematical Biosciences **136**: 1–20.

Appendix: Stochastic Sampling Algorithms for Bayesian Computation

We use a Markov chain Monte Carlo algorithm called Gibbs sampling (Gelfand and Smith 1990, Gilks et al. 1996) to draw samples from joint posterior distributions of model parameters. The Gibbs sampler is well suited to the model described in Section 2.2 because conditional posterior distributions of its parameters are relatively easy to sample. For example, when inferences are based on only 1 year of data (as in the examples of Sections 2.5.1 and 2.5.2), a sample from the joint posterior in (4) may be obtained by computing random draws from the following full-conditional distributions (modulo their normalizing constants):

$$p(\tau \mid \boldsymbol{\beta}, \rho_0, \boldsymbol{y}_1, \boldsymbol{X}_1) \ \propto \ \tau^{n/2+\varepsilon_1-1} \exp\left[-\tau\left(\varepsilon_2 + \frac{d_1(1-\rho_0^2)}{2}\right)\right] \tag{12}$$

$$p(\beta_j \mid \beta_{k(\neq j)}, \tau, \rho_0, \boldsymbol{y}_1, \boldsymbol{X}_1) \ \propto \ \exp\left[-\frac{\tau d_1(1-\rho_0^2)}{2}\right], \tag{13}$$

where $\tau = 1/\sigma^2$ and $d_1 = \sum_{i=1}^{n}(y_{i1} - \boldsymbol{x}_{i1}^T\boldsymbol{\beta})^2$. To derive these distributions, we assume a Uniform(0,1) prior for each element of $\boldsymbol{\beta}$, a Gamma$(\varepsilon_1, \varepsilon_2)$ prior for τ, and a point-mass prior for ρ (at $\rho = \rho_0$), which is not identifiable when only 1 year of data is available. Since the conditional density in (12) implies

$$\tau \mid \boldsymbol{\beta}, \rho_0, \boldsymbol{y}_1, \boldsymbol{X}_1 \sim \text{Gamma}\left(\varepsilon_1 + n/2,\ \varepsilon_2 + \frac{d_1(1-\rho_0^2)}{2}\right),$$

Gibbs samples of τ are relatively easy to compute. The conditional density in (13) does not have a familiar form, but the elements of $\boldsymbol{\beta}$ still may be sampled using an adaptive-rejection algorithm (Gilks 1992, Gilks and Wild 1992).

When inferences are based on 2 years of data (as in the example of Section 2.5.3), a sample from the joint posterior in (5) may be obtained by computing random draws from the following full-conditional distributions (modulo their normalizing constants):

$$p(\tau \mid \boldsymbol{\beta}, \rho, \boldsymbol{y}_1, \boldsymbol{y}_2, \boldsymbol{X}_1, \boldsymbol{X}_2) \ \propto \ \tau^{n+\varepsilon_1-1} \exp\Big[-\tau\Big(\varepsilon_2 + \frac{d_1(1-\rho^2)}{2} + \frac{d_2}{2}\Big)\Big], \tag{14}$$

$$p(\beta_j \mid \beta_{k(\neq j)}, \mid \tau, \rho, \boldsymbol{y}_1, \boldsymbol{y}_2, \boldsymbol{X}_1, \boldsymbol{X}_2) \;\propto\; \exp\Big[-\tau\Big(\frac{d_1(1-\rho^2)}{2} + \frac{d_2}{2}\Big)\Big], \tag{15}$$

$$p(\rho \mid \boldsymbol{\beta}, \tau, \boldsymbol{y}_1, \boldsymbol{y}_2, \boldsymbol{X}_1, \boldsymbol{X}_2) \;\propto\; (1-\rho^2)^{n/2} \exp\Big[-\tau\Big(\frac{d_1(1-\rho^2)}{2} + \frac{d_2}{2}\Big)\Big], \tag{16}$$

where $d_2 = \sum_{i=1}^n (y_{i2} - \boldsymbol{x}_{i2}^T\boldsymbol{\beta} - \rho(y_{i1} - \boldsymbol{x}_{i1}^T\boldsymbol{\beta}))^2$. To derive these distributions, we assume a Uniform(0,1) prior for each element of $\boldsymbol{\beta}$, a Gamma($\varepsilon_1, \varepsilon_2$) prior for τ, and a Uniform(-1,1) prior for ρ. As before, the conditional density in (14) has a familiar form

$$\tau \mid \boldsymbol{\beta}, \rho, \boldsymbol{y}_1, \boldsymbol{y}_2, \boldsymbol{X}_1, \boldsymbol{X}_2 \sim \text{Gamma}\left(\varepsilon_1 + n,\ \varepsilon_2 + \frac{d_1(1-\rho^2)}{2} + \frac{d_2}{2}\right)$$

that allows relatively easy sampling of τ, and samples of $\boldsymbol{\beta}$ and ρ may be computed using adaptive-rejection sampling.

Given a sample from the joint posterior distribution of model parameters, the method of composition (Tanner 1996) may be used to compute a sample from the posterior predictive distribution (7) associated with a particular set of management actions ; then, Monte Carlo integration may be used to estimate the expected loss (8) associated with this set of management actions. We demonstrate these calculations, which are rather trivial for the autoregressive model, using the example data set of Section 2.5.3. Suppose Gibbs sampling has been used to compute an arbitrarily large sample from the joint posterior distribution (5), and let $\boldsymbol{\theta}^{(r)} = (\boldsymbol{\beta}^{(r)}, \sigma^{2(r)}, \rho^{(r)})$ denote the rth element in this sample. We require a sample of the posterior predictive distribution of vegetation responses $(\tilde{\boldsymbol{y}}_3 \mid \tilde{\boldsymbol{X}}_3, \boldsymbol{y}_1, \boldsymbol{y}_2, \boldsymbol{X}_1, \boldsymbol{X}_2)$ associated with the proposed management actions specified in $\tilde{\boldsymbol{X}}_3$. By applying the method of composition to (7), the rth element $\tilde{\boldsymbol{y}}_3^{(r)}$ is easily obtained by computing a random draw from the following, n-variate normal distribution: $\mathrm{N}(\tilde{\boldsymbol{X}}_3\boldsymbol{\beta}^{(r)} + \rho^{(r)}(y_2 - \boldsymbol{X}_2\boldsymbol{\beta}^{(r)}),\ \sigma^{2(r)}I)$, where I is the $n \times n$ identity matrix. The absolute-error loss function used in the example of Section 2.5.3 is $l(\tilde{\boldsymbol{y}}_3, 0.5) = \sum_{i=1}^n |\tilde{y}_{i3} - 0.5|$. To estimate the expected loss associated with the proposed management actions $\tilde{\boldsymbol{X}}_3$, we use Monte Carlo integration to average over the posterior uncertainty expressed in the predictions of $\tilde{\boldsymbol{y}}_3$:

$$\bar{l}(\tilde{\boldsymbol{X}}_3 \mid \boldsymbol{y}_1, \boldsymbol{y}_2, \boldsymbol{X}_1, \boldsymbol{X}_2) \doteq \frac{1}{R}\sum_{r=1}^{R} l(\tilde{\boldsymbol{y}}_3^{(r)}, 0.5)$$

where R denotes the number of draws computed from the posterior predictive distribution of $\tilde{\boldsymbol{y}}_3$.

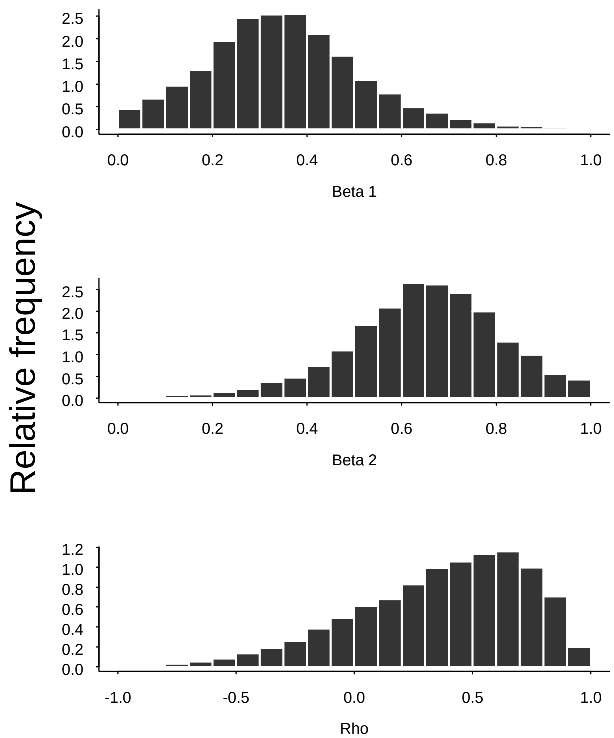

Figure 1: Histogram of the posterior distributions of β_1, β_2, and ρ estimated from the 2 years of vegetation responses given in Section 2.5.3.